

KLASIFIKASI BATIK MENGGUNAKAN METODE K-NEAREST NEIGHBOUR BERDASARKAN GRAY LEVEL CO-OCCURRENCE MATRICES (GLCM)

Hanang Wijayanto

Jurusan Teknik Informatika FIK UDINUS, Jl. Nakula No. 5-11 Semarang-50131

111201106286@mhs.dinus.ac.id

Abstrak - Batik adalah sejenis kain tertentu yang dibuat khusus dengan motif-motif yang khas yang langsung dikenali masyarakat umum. Salah satu permasalahan pada batik adalah batik memiliki motif dan warna yang sangat beragam, sehingga sulit untuk klasifikasi batik ke dalam kelas tertentu. Penelitian ini dilakukan untuk mengklasifikasikan batik ke dalam kelas-kelas berdasarkan daerah asal batik sehingga mempermudah dalam pengenalan batik dan pemahaman tentang batik. Metode yang digunakan adalah *gray level co-occurrence matrices* untuk ekstraksi ciri tekstur, sedangkan untuk menentukan kedekatan antara citra uji dengan citra latih menggunakan metode *k-nearest neighbor* berdasarkan fitur tekstur dari citra batik yang diperoleh. Fitur-fitur tekstur dicari menggunakan metode *GLCM* berdasarkan sudut 0° , 45° , 90° dan 135° . Metode evaluasi menggunakan *confusion matrix* untuk mengukur tingkat akurasi dalam proses klasifikasi. Dari hasil penelitian yang menggunakan 100 citra batik dengan 5 kelas yaitu Cirebon, Jakarta, Pekalongan, Solo dan Yogyakarta menghasilkan tingkat akurasi tertinggi pada uji coba 3 sudut 0 sebesar 57,50 % dan terendah pada uji coba 6 sudut 90 sebesar 20 %.

Kata kunci : Batik, klasifikasi, *k-nearest neighbor*, gray level co-occurrence matrix (GLCM), *confusion matrix*.

I. PENDAHULUAN

Batik adalah sejenis kain tertentu yang dibuat khusus dengan motif-motif yang khas yang langsung dikenali masyarakat umum. Batik merupakan hasil karya asli bangsa Indonesia, batik telah resmi dikukuhkan UNESCO yaitu badan dunia PBB dalam bidang kebudayaan dan pendidikan sebagai salah satu warisan dunia. Pelestarian Batik dapat dilakukan dengan berbagai cara, salah satunya dengan cara melestarikan pola-pola batik yang sangat beragam di Indonesia [1] [2].

Salah satu permasalahan dalam bidang pengenalan pola adalah klasifikasi citra ke dalam kelas tertentu. Batik dapat diklasifikasikan berdasarkan bentuk

motifnya yaitu motif geometri, motif non geometri dan motif khusus. Motif citra batik yang sangat beragam menyulitkan dalam pengenalan setiap pola citra batik. Klasifikasi data diperlukan untuk mengidentifikasi karakteristik obyek yang terkandung dalam basis data dan dikategorikan ke dalam kelompok yang berbeda. Tujuan klasifikasi batik adalah membagi citra batik ke dalam kelas-kelas motif sesuai dengan pola motifnya sehingga mudah untuk dikenali sesuai dengan cirinya. Ekstraksi ciri merupakan salah satu proses awal yang penting dalam melakukan klasifikasi citra dalam pengenalan pola. Citra batik yang terklasifikasi dengan baik akan memberikan informasi yang citra batik yang dapat digunakan untuk pelestarian

motif batik. Beberapa penelitian tentang batik telah dilakukan sebelumnya, Ade Winarni pada penelitiannya menggunakan *co-occurrence matrix* sebagai ekstraksi ciri dan *color moment* sebagai ekstraksi warna serta *k-nearest neighbor* untuk proses klasifikasi [3].

Penelitian lainnya Asri Junita Arriawati, Imam Santoso dan Yuli Christyono menggunakan matriks kookurensi sebagai ekstraksi ciri dan *k-nearest neighbor* untuk proses klasifikasi. Sedangkan Anita Ahmad Kasim dan Agus Harjoko menggunakan *gray level co-occurrence matrices* dan jaringan saraf tiruan. Pada penelitian tentang batik tersebut semuanya membahas tentang pengklasifikasian motif batik ke dalam kelas berdasarkan motifnya saja. Jadi yang diklasifikasikan dari penelitian tersebut hanya berdasarkan motif batik yang sama. Penelitian yang akan dibuat ini dilakukan untuk mengklasifikasikan batik ke dalam kelas asal daerah dari batik tersebut. Pada motif batik tersebut akan dikelompokkan ke dalam kelas-kelas asal dari batik tersebut [3] [4] [5].

II. TEORI PENUNJANG

2.1 Content Based Image Retrieval

Content Based Image Retrieval System (CBIR) merupakan suatu teknik pencarian kembali gambar yang mempunyai kemiripan karakteristik atau content dari sekumpulan gambar. Proses umum dari CBIR adalah gambar yang menjadi query dilakukan proses ekstraksi fitur, begitu halnya dengan gambar yang ada pada sekumpulan gambar juga dilakukan proses seperti pada gambar query. Fitur gambar yang dapat digunakan untuk retrieval pada system ini misalnya histogram, susunan warna, tekstur, dan shape, tipe spesifik dari obyek, tipe event tertentu, nama individu, lokasi, emosi [9].

2.2 Gray Level Co-occurrence Matrix (GLCM)

Gray Level Co-occurrence Matrix (GLCM) pertama kali diusulkan oleh Haralick pada tahun 1979 dengan 28 fitur untuk menjelaskan pola spasial. Langkah pertama untuk menghitung fitur-fitur GLCM adalah dengan mengubah citra RGB menjadi citra berskala keabuan. Langkah kedua adalah menciptakan matrik *co-occurrence* dan dilanjutkan dengan menentukan hubungan spasial antara piksel referensi dan piksel tetangga berdasarkan sudut θ dan jarak d . Langkah selanjutnya adalah menciptakan matrik simetris dengan menambahkan matrik *co-occurrence* dengan matrik transposenya. Kemudian dilakukan normalisasi terhadap matrik simetris dengan menghitung probabilitas setiap element matrik. Langkah terakhir adalah dengan menghitung fitur GLCM. Setiap fitur dihitung dengan satu piksel jarak di empat arah, yaitu 0° , 45° , 90° , dan 135° untuk mendeteksi *co-occurrence* [3] [5] [9] [11] [12] [13] [14] [15] [16] [17] [18] [19]. Terdapat 5 fitur-fitur glcm yang digunakan dalam penelitian ini, antara lain :

1. Angular Second Moment(ASM)

ASM merupakan ukuran homogenitas dari suatu citra.

$$ASM = \sum_{i=1}^L \sum_{j=1}^L (GLCM(i,j))^2$$

2. Kontras

Kontras merupakan ukuran keberadaan variasi aras keabuan piksel citra.

$$Kontras = \sum_i^L \sum_j^L |i - j|^2 GLCM(i,j)$$

3. Inverse Different Moment(IDM)

IDM digunakan untuk mengukur homogenitas.

$$IDM = \sum_{i=1}^L \sum_{j=1}^L \frac{(GLCM(i,j))^2}{1+(i-j)^2}$$

4. Entropi

Entropi menyatakan ukuran ketidakteraturan aras keabuan didalam citra.

$$\text{Entropi} = -\sum_{i=1}^L \sum_{j=1}^L (GLCM(i, j)) \log(GLCM(i, j))$$

5. Korelasi

Korelasi merupakan ukuran ketergantungan linier antar nilai aras keabuan dalam citra.

$$\text{Korelasi} = \sum_{i=1}^L \sum_{j=1}^L \frac{(i - \mu_i')(j - \mu_j')(GLCM(i, j))}{\sigma_i \sigma_j}$$

Persamaan tersebut didapat dari mean yang merupakan nilai intensitas dari citra keabuan dan standart deviasi terlebih dahulu. Standart deviasi didapat dari akar kuadrat varian yang menunjukkan sebaran nilai *piksel* dalam citra, dengan rumus sebagai berikut:

$$\begin{aligned} \text{mean } i &= \mu_i' = \sum_{i=1}^L \sum_{j=1}^L i * GLCM(i, j) \\ \text{mean } j &= \mu_j' = \sum_{i=1}^L \sum_{j=1}^L j * GLCM(i, j) \end{aligned}$$

$$\begin{aligned} \text{varian } i &= \sigma_i^2 = \sum_{i=1}^L \sum_{j=1}^L GLCM(i, j) (i - \mu_i')^2 \\ \text{varian } j &= \sigma_j^2 = \sum_{i=1}^L \sum_{j=1}^L GLCM(i, j) (j - \mu_j')^2 \end{aligned}$$

$$\text{standart deviasi } i = \sigma_i = \sqrt{\sigma_i^2}$$

$$\text{standart deviasi } j = \sigma_j = \sqrt{\sigma_j^2}$$

2.3 K-Nearest Neighbour

K-Nearest Neighbor (KNN) adalah suatu metode yang menggunakan algoritma *supervised* dimana hasil dari *query instance* yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada KNN. Tujuan dari algoritma ini adalah mengklasifikasikan obyek baru berdasarkan atribut dan *training sample*. Algoritma metode KNN sangatlah sederhana, bekerja berdasarkan jarak terpendek dari *query instance* ke *training sample* untuk menentukan KNN-nya. *Training sample* diproyeksikan ke ruang berdimensi banyak, dimana masing-

masing dimensi merepresentasikan fitur dari data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi *training sample*. Sebuah titik pada ruang ini ditandai kelas *c* jika kelas *c* merupakan klasifikasi yang paling banyak ditemui pada *k* buah tetangga terdekat dari titik tersebut. Dekat atau jauhnya tetangga biasanya dihitung berdasarkan *Euclidean Distance* yang direpresentasikan sebagai berikut [5] [20]:

$$d = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2}$$

$$d = \sqrt{\sum_{i=1}^n (a_i - b_i)^2}$$

2.4 Confusion matrix

Confusion matrix merupakan sebuah tabel yang terdiri dari atas banyaknya baris data uji yang diprediksi benar dan tidak benar oleh model klasifikasi, tabel ini diperlukan untuk menentukan kinerja suatu model klasifikasi [21] [22].

Tabel 1. Tabel Confusion Matrix

		Predicted Class	
Actual Class	Class	Class=1	Class=0
	Class=1	F 11	F 10
	Class=0	F 01	F 00

Perhitungan akurasi dengan menggunakan tabel *confusion matrix* adalah sebagai berikut :

$$\text{akurasi} = \frac{F 11 + F 00}{F 11 + F 10 + F 01 + F 00}$$

III. HASIL &IMPLEMENTASI

Hasil pengujian dalam klasifikasi batik dengan 6 kali percobaan dengan 5 kelas yaitu Cirebon, Jakarta, Pekalongan, Solo dan Yogyakarta menunjukkan hasil yang berbeda-beda untuk ke-empat sudut θ yang digunakan, ada beberapa data yang

dapat terklasifikasi dengan benar sesuai dengan kelas dan ada pula yang tidak terklasifikasi dengan benar. Hasil data dari proses klasifikasi yang diikuti untuk ke-empat sudut yang berbeda tersebut antara lain sebagai berikut:

Tabel 2 Tingkat Akurasi

	1	2	3	4	5	6
0	41.67%	52%	57.50%	53.33%	40%	40%
45	31.67%	44%	55%	53.33%	45%	40%
90	28.33%	36%	42.50%	40%	40%	20%
135	26.67%	42%	37.50%	43.33%	35%	30%

Dari perhitungan diatas, berdasarkan tabel 2 dengan 6 kali uji coba diperoleh akurasi tertinggi pada uji coba 3 sudut 0 yaitu 57,50 % dan terendah pada uji coba 6 sudut 90 yaitu 20 %.

IV. PENUTUP

Berdasarkan penelitian yang telah dilakukan oleh peneliti pada klasifikasi batik berdasarkan daerah asal menghasilkan klasifikasi batik dengan algoritma *k-nearest neighbour* dapat mengklasifikasikan batik berdasarkan kelas - kelas yang ada. Namun tidak semua batik dapat terklasifikasikan dengan benar. Itu ditunjukan dengan tingkat akurasi yang dihasilkan dalam proses klasifikasi belum terlalu baik. Berdasarkan uji coba yang telah dilakukan menggunakan jumlah data *training* dan *testing* yang berbeda-beda menghasilkan tingkat akurasi tertinggi pada uji coba 3 sudut 0 yaitu 57,50 % dan tingkat akurasi terendah pada uji coba 6 sudut 90 yaitu 20 %.

DAFTAR PUSTAKA

[1] Casta and Taruna, Batik Cirebon, Warisan Budaya Dunia Asli Indonesia, Cirebon: Badan

Komunikasi, Kebudayaan dan Pariwisata Kabupaten Cirebon, 2007.

- [2] A. Wulandari, Batik Nusantara Makna Filosofi, Cara Pembuatan, dan Industri Batik, Yogyakarta: ANDI, 2011.
- [3] A. A. Kasim and A. Harjoko, "Klasifikasi Citra Batik Menggunakan Jaringan Syaraf Tiruan Berdasarkan Gray Level Co-Occurrence Matrices (GLCM)," *Seminar Nasional Aplikasi Teknologi Informasi (SNATI)*, 21 Juni 2014.
- [4] A. Winarni, "Ekstraksi Ciri Citra Batik Menggunakan Co-Occurrence Matrix, Color Moment dan K-Nearest Neighbor," *LONTAR KOMPUTER*, vol. 4, no. 1, April 2013.
- [5] A. J. Arriawati, I. Santoso and Y. Christyono, "Klasifikasi Citra Tekstur Menggunakan K-Nearest Neighbour Berdasarkan Ekstraksi Ciri Metode Matriks Kookurensi," Universitas Diponegoro, Semarang.
- [6] E. W. Wardani, "Pengenalan Motif Batik Menggunakan Metode Transformasi Paket Wavelet," Universitas Widyatama, Bandung, 2013.
- [7] N. Syafitri, "Perbandingan Metode K-Nearest Neighbor (KNN) Dan Metode Nearest Cluster Classifier (NCC) Dalam Pengklasifikasian Kualitas Batik Tulis," *JURNAL TEKNOLOGI INFORMASI & PENDIDIKAN*, vol. 2, no. 1, September 2010.
- [8] A. Kurniawardhani, N. Suciati and I. Arieshanti, "Klasifikasi Citra Batik

- Menggunakan Metode Ekstraksi Ciri Yang Invariant Terhadap Rotasi," *JUTI*, vol. 12, no. 2, pp. 48 - 60, Juli 2014.
- [9] A. Kadir and A. Susanto, Pengolahan Citra Teori dan Aplikasi, Yogyakarta, 2012.
- [10] Z. Budiarto, "Identifikasi Macan Tutul Dengan Metode Grey Level Coocurent Matrix (GLCM)," Univesitas Stikubank, Semarang.
- [11] A. E. Minarno and N. Suciati, "Batik Image Retrieval Based on Color Difference Histogram and Gray Level Co-Occurrence Matrix," *TELKOMNIKA*, vol. 12, no. 3, pp. 597-604, September 2014.
- [12] A. A. Pratama, N. Suciati and D. Purwitasari, "Implementasi Fuzzy C-Means untuk Pengelompokan Citra Batik Berdasarkan Motif dengan Fitur Tekstur," *JURNAL TEKNIK POMITS*, vol. 1, no. 1, pp. 1-4, 2012.
- [13] B. S. V, A. Unnikrishnan and K. Balakrishnan, "Grey Level Co-Occurrence Matrices:Generalisation And Some New Features," *International Journal of Computer Science, Engineering and Information Technology (IJCSEIT)*, vol. 2, no. 2, April 2012.
- [14] S. K and M. L, "An Efficient Image Retrieval Based on Color, Texture (GLCM & CCM) features, and Genetic-Algorithm," *International Journal Of Merging TechnologyAnd Advanced Research In Computing*, pp. 1-9.
- [15] P. Maheshwary and N. Sricastava, "Prototype System for Retrieval of Remote Sensing Images based on Color Moment and Gray Level Co-Occurrence Matrix," *IJCSI International Journal of Computer Science Issues*, vol. 3, pp. 20-23, 2009.
- [16] R. B and K. R. Chandran, "Content Based Medical Image Retrieval with Texture Content Using Gray Level Co-occurence Matrix and K-Means Clustering Algorithm," *Journal of Computer Science*, vol. 8, no. 7, pp. 1070-1076, 2012.
- [17] M.-W. Lin, J.-R. Tapamo and B. Ndovie, "A Texture-based Method for Document Segmentation and Classification," *ARIMA/SACJ*, no. 36, pp. 49-56, 2006.
- [18] F. Albrechtsen, "Statistical Texture Measures Computed from Gray Level Coocurrence Matrices," *Image Processing Laboratory Department of Informatics University of Oslo*, pp. 1-14, 5 November 2008.
- [19] E. K. Ratnasari, R. H. Ginardi and C. Fatichah, "Pengenalan Penyakit Noda Pada Citra DaunTebu Berdasarkan Ciri Tekstur Fractal Dimension Co-Occurrence Matrix Dan L*a*b* Color Moments," *JUTI*, vol. 12, no. 2, pp. 27-36, Juli 2014.
- [20] M. I. Sikki, "Pengenalan Wajah Menggunakan K-Nearest Neighbour Dengan Praproses Transformasi Wavelet," *Jurnal Paradigma*, vol. 10, no. 2, Desember 2009.

- [21] A. Andriani, "Penerapan Algoritma C4.5 pada Program Klasifikasi Mahasiswa Dropout," *Seminar Nasional Matematika*, 2012.
- [22] I. Juniawan, "Klasifikasi Dokumen Teks Berbahasa Menggunakan Minor Component Analysis," Institut Pertanian Bogor, Bogor, 2009.